



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년01월26일

(11) 등록번호 10-1480668

(24) 등록일자 2015년01월02일

(51) 국제특허분류(Int. Cl.)

H04B 1/40 (2015.01) G06F 3/01 (2006.01)

G10L 15/00 (2006.01)

(21) 출원번호 10-2014-0033566

(22) 출원일자 2014년03월21일

심사청구일자 2014년03월21일

(56) 선행기술조사문헌

JP2006106711 A

(73) 특허권자

충남대학교산학협력단

대전광역시 유성구 대학로 99 (궁동, 충남대학교)

(72) 발명자

손진훈

대전 유성구 가정로 43, 107동 904호 (신성동, 삼성한울아파트)

박정식

대전광역시 중구 충무로 107번길 100 (대흥동, 센트럴자이) 113동 1502호

박지은

대전 유성구 엑스포로339번길 320, 7-403 (원촌동, 싸이언스빌)

(74) 대리인

특허법인 아이퍼스

전체 청구항 수 : 총 11 항

심사관 : 임동우

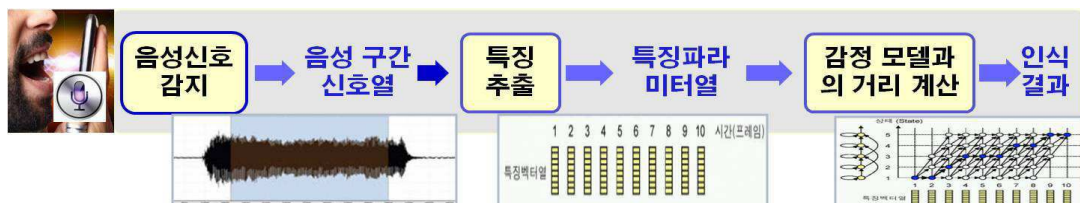
(54) 발명의 명칭 음성을 이용한 감정 인식 어플리케이션이 구비된 단말기 및 그 제어방법

(57) 요약

본 발명은 사용자의 음성을 이용하여 사용자의 감정을 인식하고, 사용자의 감정 상태에 따른 적절한 반응을 출력할 수 있는 단말기 및 그 제어방법에 관한 것이다. 본 발명의 일례와 관련된 단말기는, 사용자의 음성을 포함하는 제 1 음성을 획득하기 위한 마이크, 복수의 감정과 관련된 음향 모델 데이터를 저장하는 메모리, 상기 제 1

(뒷면에 계속)

대표도 - 도8



음성을 수신받고 상기 음향 모델 데이터와 상기 제 1 음성을 이용하여 상기 복수의 감정 중 상기 사용자의 감정을 인식할 수 있는 감정 인식 수단, 시각정보를 표시하는 디스플레이부, 청각정보를 출력하는 음향 출력 모듈, 진동을 발생시키는 햅틱 모듈 및 후각정보를 발생시키는 방향부로 구성된 출력부 및 상기 사용자로부터 상기 단말기에 내장 설치된 감정 인식 어플리케이션의 동작에 관한 명령을 입력받기 위한 입력부를 포함하되, 상기 음향 모델 데이터는 상기 복수의 감정 각각에 대응하는 상기 음성의 특징 파라미터에 대한 정보이고, 상기 사용자가 상기 입력부를 이용하여 상기 감정 인식 어플리케이션의 실행 명령을 입력하는 경우, 상기 감정 인식 수단과 관련된 소정의 실행창이 상기 디스플레이부에 표시되어 상기 제 1 음성의 녹음을 유도하고, 상기 출력부는 상기 제 1 음성에 기초하여 인식된 상기 사용자의 감정에 대응하는 반응을 출력할 수 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호	2013K000332
부처명	미래창조과학부
연구관리전문기관	한국표준연구원
연구사업명	신기술융합형 성장동력사업
연구과제명	인간-기계 인터페이스를 위한 인지-감정 통합모델 개발
기 여 율	1/1
주관기관	한국표준연구원
연구기간	2013.07.01 ~ 2014.06.30

특허청구의 범위

청구항 1

사용자의 음성을 이용하여 기 설정된 기준에 따라 분류된 복수의 감정을 인식할 수 있는 단말기에 있어서,

상기 사용자의 음성을 포함하는 제 1 음성을 획득하기 위한 마이크;

상기 복수의 감정과 관련된 음향 모델 데이터를 저장하는 메모리;

상기 제 1 음성을 수신받고, 상기 음향 모델 데이터와 상기 제 1 음성을 이용하여 상기 복수의 감정 중 상기 사용자의 감정을 인식할 수 있는 감정 인식 수단;

시각정보를 표시하는 디스플레이부, 청각정보를 출력하는 음향 출력 모듈, 진동을 발생시키는 햅틱 모듈 및 후각정보를 발생시키는 방향부로 구성된 출력부; 및

상기 사용자로부터 상기 단말기에 내장 설치된 감정 인식 어플리케이션의 동작에 관한 명령을 입력받기 위한 입력부;를 포함하되,

상기 음향 모델 데이터는 상기 복수의 감정 각각에 대응하는 상기 음성의 특징 파라미터에 대한 정보이고,

상기 사용자가 상기 입력부를 이용하여 상기 감정 인식 어플리케이션의 실행 명령을 입력하는 경우, 상기 감정 인식 수단과 관련된 소정의 실행창이 상기 디스플레이부에 표시되어 상기 제 1 음성의 녹음을 유도하고, 상기 출력부는 상기 제 1 음성에 기초하여 인식된 상기 사용자의 감정에 대응하는 반응을 출력하며,

상기 감정 인식 수단은,

상기 제 1 음성의 신호 에너지(signal energy)에 기초하여 상기 제 1 음성에 녹음된 상기 사용자의 음성을 감지하는 감지부;

상기 감지된 상기 사용자의 음성에서 제 1 특징 파라미터를 추출하는 추출부; 및

상기 제 1 특징 파라미터를 상기 음향 모델 데이터와 비교하여 상기 사용자의 감정을 판단하는 인식부;를 더 포함하고,

상기 인식부는, 상기 메모리에 저장된 상기 음향 모델 데이터와 상기 제 1 특징 파라미터의 관측 확률값을 계산하고, 상기 계산된 관측 확률값 중 가장 높은 관측 확률값을 갖는 음향 모델 데이터에 대응하는 감정을 상기 사용자의 감정으로 판단하며,

상기 메모리에 저장된 음향 모델 데이터는, 은닉 마르코프 모델(Hidden Markov Model)에 기초하여 생성되고,

상기 사용자가 상기 입력부를 이용하여 상기 실행창에 표시된 업데이트 실행 명령을 입력하는 경우, 상기 사용자에게 대한 음향 모델 데이터가 상기 은닉 마르코프 모델에 기초하여 생성되어 상기 메모리에 추가적으로 저장되는 것을 특징으로 하는 단말기.

청구항 2

제 1항에 있어서,

상기 반응은,

상기 디스플레이부에서의 제 1 자료 표시, 상기 음향 출력 모듈에서의 음향 출력, 상기 햅틱 모듈에서 소정 세기와 패턴을 갖는 진동의 발생 및 상기 방향부에서 발생하는 향기 중 적어도 하나를 포함하고,

상기 제 1 자료는 기 설정된 텍스트, 배경화면 및 이모티콘 중 적어도 하나인 것을 특징으로 하는 단말기.

청구항 3

제 1항에 있어서,

상기 실행창은 상기 사용자로부터 녹음 정보를 입력받아 상기 제 1 음성의 녹음에 이용하고,

상기 녹음 정보는, 상기 사용자의 성별, 상기 제 1 음성의 녹음 시작 시점 및 상기 제 1 음성이 녹음되는 시간 중 적어도 하나에 대한 정보를 포함하는 것을 특징으로 하는 단말기.

청구항 4

제 3항에 있어서,

상기 메모리에는,

상기 성별에 따라 서로 다른 음향 모델 데이터가 저장된 것을 특징으로 하는 단말기.

청구항 5

삭제

청구항 6

제 1항에 있어서,

상기 추출부에 의하여 추출된 상기 특징 파라미터는 MFCC(Mel-Frequency Cepstral Coefficient) 및 로그 에너지(Log-energy)를 포함하는 것을 특징으로 하는 단말기.

청구항 7

제 6항에 있어서,

상기 MFCC는,

상기 감지부에 의하여 감지된 사용자의 음성에 FFT(Fast Fourier Transform)를 적용하고, 상기 FFT가 적용된 결과에 Mel-scale의 필터 뱅크를 적용하여 파워스펙트럼을 구하며, 상기 파워스펙트럼에 로그(log)를 취하고, 상기 로그를 취한 결과에 DCT(Discrete Cosine Transform)를 적용하여 얻어지는 것을 특징으로 하는 단말기.

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

제 1항에 있어서,

상기 기 설정된 기준에 따라 분류된 복수의 감정은,

중립, 기쁨, 분노 및 슬픔 중 적어도 하나를 포함하는 것을 특징으로 하는 단말기.

청구항 12

제 1항에 있어서,

상기 사용자가 상기 단말기를 이용하여 상대방과 통화 중인 경우,

상기 단말기는 상기 감정 인식 수단을 이용하여 상기 상대방의 감정을 인식하고,

상기 출력부는 상기 인식된 상대방의 감정에 대응하여 반응을 출력하는 것을 특징으로 하는 단말기.

청구항 13

사용자의 음성을 이용하여 기 설정된 기준에 따라 분류된 복수의 감정을 인식할 수 있는 단말기를 제어하는 방

법에 있어서,

상기 사용자로부터 상기 단말기에 내장 설치된 감정 인식 어플리케이션의 실행 명령을 입력받는 제 1 단계;

감정 인식 수단과 관련된 소정의 실행창이 디스플레이부에 표시되는 제 2 단계;

상기 사용자가 상기 사용자의 음성을 포함하는 제 1 음성에 대한 녹음 정보를 입력하는 제 3 단계;

상기 제 1 음성이 녹음되는 제 4 단계;

상기 감정 인식 수단으로 상기 제 1 음성이 전송되는 제 5 단계;

상기 감정 인식 수단이 메모리에 저장된 음향 모델 데이터와 상기 제 1 음성을 이용하여 상기 복수의 감정 중 상기 사용자의 감정을 인식하는 제 6 단계; 및

출력부가 상기 제 1 음성에 기초하여 인식된 상기 사용자의 감정에 대응하는 반응을 출력하는 제 7 단계;를 포함하되,

상기 음향 모델 데이터는 상기 복수의 감정 각각에 대응하는 상기 음성의 특징 파라미터에 대한 정보이고,

상기 출력부는 시각정보를 표시하는 상기 디스플레이부, 청각정보를 출력하는 음향 출력 모듈, 진동을 발생시키는 햅틱 모듈 및 후각정보를 발생시키는 방향부로 구성되며,

상기 제 6 단계는,

상기 제 1 음성의 신호 에너지(signal energy)에 기초하여 상기 제 1 음성에 녹음된 상기 사용자의 음성을 감지하는 제 6-1 단계;

상기 감지된 상기 사용자의 음성에서 제 1 특징 파라미터를 추출하는 제 6-2 단계; 및

상기 제 1 특징 파라미터를 상기 음향 모델 데이터와 비교하여 상기 사용자의 감정을 판단하는 제 6-3 단계;를 더 포함하고,

상기 제 6-3 단계에서는, 상기 메모리에 저장된 상기 음향 모델 데이터와 상기 제 1 특징 파라미터의 관측 확률값을 계산하고, 상기 계산된 관측 확률값 중 가장 높은 관측 확률값을 갖는 음향 모델 데이터에 대응하는 감정을 상기 사용자의 감정으로 판단하며,

상기 메모리에 저장된 음향 모델 데이터는, 은닉 마르코프 모델(Hidden Markov Model)에 기초하여 생성되고,

상기 사용자가 상기 입력부를 이용하여 상기 실행창에 표시된 업데이트 실행 명령을 입력하는 경우, 상기 사용자에게 대한 음향 모델 데이터가 상기 은닉 마르코프 모델에 기초하여 생성되어 상기 메모리에 추가적으로 저장되는 것을 특징으로 하는 단말기의 제어방법.

청구항 14

제 13항에 있어서,

상기 반응은,

상기 디스플레이부에서의 제 1 자료 표시, 상기 음향 출력 모듈에서의 음향 출력, 상기 햅틱 모듈에서 소정 세기와 패턴을 갖는 진동의 발생 및 상기 방향부에서 발생하는 향기 중 적어도 하나를 포함하고,

상기 제 1 자료는 기 설정된 텍스트, 배경화면 및 이모티콘 중 적어도 하나인 것을 특징으로 하는 단말기의 제어방법.

청구항 15

삭제

청구항 16

제 13항에 있어서,

상기 기 설정된 기준에 따라 분류된 복수의 감정은,

중립, 기쁨, 분노 및 슬픔 중 적어도 하나를 포함하는 것을 특징으로 하는 단말기의 제어방법.

명세서

기술분야

- [0001] 본 발명은 감정 인식 어플리케이션이 구비된 단말기 및 그 제어방법에 관한 것으로서, 보다 상세하게는 사용자의 음성을 이용하여 사용자의 감정을 인식하고, 사용자의 감정 상태에 따른 적절한 반응을 출력할 수 있는 단말기 및 그 제어방법에 관한 것이다.

배경기술

- [0002] 멀티모달 감정 인식(multimodal emotion recognition)은 facial expression, speech, gesture, gaze, head movement, context 등의 여러 가지 정보를 이용해서 감정을 파악하는 것을 의미하고, 멀티모달 인터페이스를 통한 멀티 모달 정보가 입력되면 각각의 모달리티에서 입력된 정보가 융합되어 분석될 수 있다. 또한, 멀티 모달리티에서 입력된 정보의 특징을 추출하고 분류하기 위해서는 다양한 학습 알고리즘이 사용될 수 있다. 이때, 어떠한 학습 알고리즘이 적용되는지 여부에 따라 결과를 분석하여 인식하는 어려움이 달라질 수 있다.
- [0003] 사용자의 감정을 인식할 수 있는 기능은 지능형 인터페이스의 중요 요소인데, 이를 위하여 사용자의 얼굴 표정, 음성 등의 특징을 이용한 감정인식 기술이 시도되고 있다.
- [0004] 그러나, 음성에 기반한 감정인식 기술의 경우 대부분 그 인식률이 만족할 만한 수준에 도달하지 못하는 문제점이 있었다.
- [0005] 또한, 자연스러운 감정을 유발하여 객체반응을 측정하지 않고, 인위적인 감정상태의 객체반응 데이터를 사용함으로써, 실제 이벤트와 매칭되지 않는 경우가 빈번하게 발생 된다는 문제점도 있었다.
- [0006] 이에 따라, 사용자의 음성을 이용하여 사용자의 감정을 보다 정확하게 인식할 수 있으며, 상기 사용자의 감정상태를 고려하여 적절한 피드백을 제공할 수 있는 감정 인식 어플리케이션이 구비된 단말기의 개발이 요구되고 있는 실정이다.

선행기술문헌

특허문헌

- [0007] (특허문헌 0001) 대한민국 공개특허 제10-2002-0018541호

비특허문헌

- [0008] (비특허문헌 0001) F. Samaria and S. Young, "HMM based architecture for face identification", Image and Computer Vision, vol. 12, pp.537-543, October 1994

발명의 내용

해결하려는 과제

- [0009] 본 발명은 상기와 같은 종래의 문제점을 해결하기 위하여 안출된 것으로서, 사용자의 음성을 이용하여 사용자의 감정을 인식하고, 사용자의 감정 상태에 따른 적절한 반응을 출력할 수 있는 단말기 및 그 제어방법을 사용자에게 제공하는 데 그 목적이 있다.
- [0010] 구체적으로, 본 발명은 사용자가 느끼는 복수의 감정에 따른 음성의 특징 파라미터를 활용하여 사용자의 감정을 보다 정확하게 인식할 수 있는 단말기 및 그 제어방법을 사용자에게 제공하는 데 그 목적이 있다.
- [0011] 또한, 본 발명은 사용자가 느끼는 감정에 대응하는 피드백을 단말기를 이용하여 출력하여 사용자의 감정상태를 고려한 피드백의 제공이 가능한 단말기 및 그 제어방법을 사용자에게 제공하는 데 그 목적이 있다.

[0012] 한편, 본 발명에서 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급하지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0013] 사용자의 음성을 이용하여 기 설정된 기준에 따라 분류된 복수의 감정을 인식할 수 있는 단말기에 있어서, 상술한 과제를 실현하기 위한 본 발명의 일례와 관련된 단말기는, 상기 사용자의 음성을 포함하는 제 1 음성을 획득하기 위한 마이크, 상기 복수의 감정과 관련된 음향 모델 데이터를 저장하는 메모리, 상기 제 1 음성을 수신받고 상기 음향 모델 데이터와 상기 제 1 음성을 이용하여 상기 복수의 감정 중 상기 사용자의 감정을 인식할 수 있는 감정 인식 수단, 시각정보를 표시하는 디스플레이부, 청각정보를 출력하는 음향 출력 모듈, 진동을 발생시키는 햅틱 모듈 및 후각정보를 발생시키는 방향부로 구성된 출력부 및 상기 사용자로부터 상기 단말기에 내장 설치된 감정 인식 어플리케이션의 동작에 관한 명령을 입력받기 위한 입력부를 포함하되, 상기 음향 모델 데이터는 상기 복수의 감정 각각에 대응하는 상기 음성의 특징 파라미터에 대한 정보이고, 상기 사용자가 상기 입력부를 이용하여 상기 감정 인식 어플리케이션의 실행 명령을 입력하는 경우, 상기 감정 인식 수단과 관련된 소정의 실행창이 상기 디스플레이부에 표시되어 상기 제 1 음성의 녹음을 유도하고, 상기 출력부는 상기 제 1 음성에 기초하여 인식된 상기 사용자의 감정에 대응하는 반응을 출력할 수 있다.

[0014] 또한, 상기 반응은, 상기 디스플레이부에서의 제 1 자료 표시, 상기 음향 출력 모듈에서의 음향 출력, 상기 햅틱 모듈에서 소정 세기와 패턴을 갖는 진동의 발생 및 상기 방향부에서 발생하는 향기 중 적어도 하나를 포함할 수 있고, 상기 제 1 자료는 기 설정된 텍스트, 배경화면 및 이모티콘 중 적어도 하나이다.

[0015] 또한, 상기 실행창은 상기 사용자로부터 녹음 정보를 입력받아 상기 제 1 음성의 녹음에 이용하고, 상기 녹음 정보는, 상기 사용자의 성별, 상기 제 1 음성의 녹음 시작 시점 및 상기 제 1 음성이 녹음되는 시간 중 적어도 하나에 대한 정보를 포함할 수 있다.

[0016] 또한, 상기 메모리에는, 상기 성별에 따라 서로 다른 음향 모델 데이터가 저장될 수 있다.

[0017] 또한, 상기 감정 인식 수단은, 상기 제 1 음성의 신호 에너지(signal energy)에 기초하여 상기 제 1 음성에 녹음된 상기 사용자의 음성을 감지하는 감지부, 상기 감지된 상기 사용자의 음성에서 제 1 특징 파라미터를 추출하는 추출부 및 상기 제 1 특징 파라미터를 상기 음향 모델 데이터와 비교하여 상기 사용자의 감정을 판단하는 인식부를 더 포함할 수 있다.

[0018] 또한, 상기 추출부에 의하여 추출된 상기 특징 파라미터는 MFCC(Mel-Frequency Cepstral Coefficient) 및 로그 에너지(Log-energy)를 포함할 수 있다.

[0019] 또한, 상기 MFCC는, 상기 감지부에 의하여 감지된 사용자의 음성에 FFT(Fast Fourier Transform)를 적용하고, 상기 FFT가 적용된 결과에 Mel-scale의 필터 뱅크를 적용하여 파워스펙트럼을 구하며, 상기 파워스펙트럼에 로그(log)를 취하고, 상기 로그를 취한 결과에 DCT(Discrete Cosine Transform)를 적용하여 얻어질 수 있다.

[0020] 또한, 상기 인식부는, 상기 메모리에 저장된 상기 음향 모델 데이터와 상기 제 1 특징 파라미터의 관측 확률값을 계산하고, 상기 계산된 관측 확률값 중 가장 높은 관측 확률값을 갖는 음향 모델 데이터에 대응하는 감정을 상기 사용자의 감정으로 판단할 수 있다.

[0021] 또한, 상기 메모리에 저장된 음향 모델 데이터는, 은닉 마르코프 모델(Hidden Markov Model)에 기초하여 생성될 수 있다.

[0022] 또한, 상기 사용자가 상기 입력부를 이용하여 상기 실행창에 표시된 업데이트 실행 명령을 입력하는 경우, 상기 사용자에게 대한 음향 모델 데이터가 상기 은닉 마르코프 모델에 기초하여 생성되어 상기 메모리에 추가적으로 저장될 수 있다.

[0023] 또한, 상기 기 설정된 기준에 따라 분류된 복수의 감정은, 중립, 기쁨, 분노 및 슬픔 중 적어도 하나를 포함할 수 있다.

[0024] 또한, 상기 사용자가 상기 단말기를 이용하여 상대방과 통화 중인 경우, 상기 단말기는 상기 감정 인식 수단을 이용하여 상기 상대방의 감정을 인식하고, 상기 출력부는 상기 인식된 상대방의 감정에 대응하여 반응을 출력할 수 있다.

- [0025] 사용자의 음성을 이용하여 기 설정된 기준에 따라 분류된 복수의 감정을 인식할 수 있는 단말기를 제어하는 방법에 있어서, 상술한 과제를 실현하기 위한 본 발명의 일례와 관련된 단말기의 제어방법은, 상기 사용자로부터 상기 단말기에 내장 설치된 감정 인식 어플리케이션의 실행 명령을 입력받는 제 1 단계, 감정 인식 수단과 관련된 소정의 실행창이 디스플레이부에 표시되는 제 2 단계, 상기 사용자가 상기 사용자의 음성을 포함하는 제 1 음성에 대한 녹음 정보를 입력하는 제 3 단계, 상기 제 1 음성이 녹음되는 제 4 단계, 상기 감정 인식 수단으로 상기 제 1 음성이 전송되는 제 5 단계, 상기 감정 인식 수단이 음향 모델 데이터와 상기 제 1 음성을 이용하여 상기 복수의 감정 중 상기 사용자의 감정을 인식하는 제 6 단계 및 출력부가 상기 제 1 음성에 기초하여 인식된 상기 사용자의 감정에 대응하는 반응을 출력하는 제 7 단계를 포함하되, 상기 음향 모델 데이터는 상기 복수의 감정 각각에 대응하는 상기 음성의 특징 파라미터에 대한 정보이고, 상기 출력부는 시각정보를 표시하는 상기 디스플레이부, 청각정보를 출력하는 음향 출력 모듈, 진동을 발생시키는 햅틱 모듈 및 후각정보를 발생시키는 방향부로 구성될 수 있다.
- [0026] 또한, 상기 반응은, 상기 디스플레이부에서의 제 1 자료 표시, 상기 음향 출력 모듈에서의 음향 출력, 상기 햅틱 모듈에서 소정 세기와 패턴을 갖는 진동의 발생 및 상기 방향부에서 발생하는 향기 중 적어도 하나를 포함할 수 있고, 상기 제 1 자료는 기 설정된 텍스트, 배경화면 및 이모티콘 중 적어도 하나이다.
- [0027] 또한, 상기 제 6 단계는, 상기 제 1 음성의 신호 에너지(signal energy)에 기초하여 상기 제 1 음성에 녹음된 상기 사용자의 음성을 감지하는 단계, 상기 감지된 상기 사용자의 음성에서 제 1 특징 파라미터를 추출하는 단계 및 상기 제 1 특징 파라미터를 상기 음향 모델 데이터와 비교하여 상기 사용자의 감정을 판단하는 단계를 더 포함할 수 있다.
- [0028] 또한, 상기 기 설정된 기준에 따라 분류된 복수의 감정은, 중립, 기쁨, 분노 및 슬픔 중 적어도 하나를 포함할 수 있다.

발명의 효과

- [0029] 본 발명은 사용자의 음성을 이용하여 사용자의 감정을 인식하고, 사용자의 감정 상태에 따른 적절한 반응을 출력할 수 있는 단말기 및 그 제어방법을 사용자에게 제공할 수 있다.
- [0030] 구체적으로, 본 발명은 사용자가 느끼는 복수의 감정에 따른 음성의 특징 파라미터를 활용하여 사용자의 감정을 보다 정확하게 인식할 수 있는 단말기 및 그 제어방법을 사용자에게 제공할 수 있다.
- [0031] 또한, 본 발명은 사용자가 느끼는 감정에 대응하는 피드백을 단말기를 이용하여 출력하여 사용자의 감정상태를 고려한 피드백의 제공이 가능한 단말기 및 그 제어방법을 사용자에게 제공할 수 있다.
- [0032] 한편, 본 발명에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0033] 본 명세서에 첨부되는 다음의 도면들은 본 발명의 바람직한 일 실시예를 예시하는 것이며, 발명의 상세한 설명과 함께 본 발명의 기술적 사상을 더욱 이해시키는 역할을 하는 것이므로, 본 발명은 그러한 도면에 기재된 사항에만 한정되어 해석되어서는 아니 된다.
- 도 1은 본 발명의 감정 인식 어플리케이션이 구비된 단말기의 블록 구성도의 일례를 나타낸 것이다.
- 도 2a 및 도 2b는 본 발명과 관련된 단말기의 전면 사시도 및 후면 사시도의 일례를 나타낸다.
- 도 3은 본 발명과 관련된 단말기의 일 작동 상태를 설명하기 위한 단말기의 정면도이다.
- 도 4는 본 발명에 따른 감정 인식 어플리케이션의 실행에 의한 실행창의 일 실시예를 나타낸다.
- 도 5a 및 도 5b는 감정 인식 수단의 추출부가 특징 파라미터를 추출하는 과정의 일 실시예를 개략적으로 나타낸 것이다.
- 도 6a 및 도 6b는 감정 인식 수단의 인식부가 감정을 훈련 및 인식하는 과정의 일 실시예를 개략적으로 나타낸 것이다.

도 7은 본 발명의 감정 인식 수단에 따른 인식률 및 인식시간에 대한 실험 결과를 나타낸다.

도 8은 본 발명의 감정 인식 수단이 사용자의 감정을 인식하는 과정을 개략적으로 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

- [0034] 이하, 도면을 참조하여 본 발명의 바람직한 일 실시예에 대해서 설명한다. 또한, 이하에 설명하는 일 실시예는 특허청구범위에 기재된 본 발명의 내용을 부당하게 한정하지 않으며, 본 실시 형태에서 설명되는 구성 전체가 본 발명의 해결 수단으로서 필수적이라고는 할 수 없다.
- [0035] 이하에서는, 본 발명이 제안하고자 하는 음성을 이용한 감정 인식 어플리케이션이 구비된 단말기를 구체적으로 설명한다.
- [0036] 본 명세서에서 설명되는 이동 단말기, 이동단말, 단말 또는 휴대 단말기에는 휴대폰, 스마트 폰(smart phone), 노트북 컴퓨터(laptop computer), 디지털방송용 단말기, PDA(Personal Digital Assistants), PMP(Portable Multimedia Player), 네비게이션 등이 포함될 수 있다. 그러나, 본 명세서에 기재된 실시예에 따른 구성은 휴대 단말기에만 적용 가능한 경우를 제외하면, 디지털 TV, 데스크탑 컴퓨터 등과 같은 고정 단말기에도 적용될 수도 있음을 본 기술분야의 당업자라면 쉽게 알 수 있을 것이다.
- [0037] 도 1은 본 발명의 감정 인식 어플리케이션이 구비된 단말기의 블록 구성도의 일례를 나타낸 것이다.
- [0038] 상기 단말기(1100)는 무선 통신부(1110), A/V(Audio/Video) 입력부(1120), 입력부(1130), 센싱부(1140), 출력부(1150), 메모리(1160), 인터페이스부(1170), 제어부(1180) 및 전원 공급부(1190) 등을 포함할 수 있다. 도 1에 도시된 구성요소들이 필수적인 것은 아니어서, 그보다 많은 구성요소들을 갖거나 그보다 적은 구성요소들을 갖는 단말기가 구현될 수도 있다.
- [0039] 이하, 상기 구성요소들에 대해 차례로 살펴본다.
- [0040] 무선 통신부(1110)는 단말기(1100)와 무선 통신 시스템 사이 또는 단말기(1100)와 단말기(1100)가 위치한 네트워크 사이의 무선 통신을 가능하게 하는 하나 이상의 모듈을 포함할 수 있다. 예를 들어, 무선 통신부(1110)는 방송 수신 모듈(1111), 이동통신 모듈(1112), 무선 인터넷 모듈(1113), 근거리 통신 모듈(1114) 및 위치정보 모듈(1115) 등을 포함할 수 있다.
- [0041] 방송 수신 모듈(1111)은 방송 채널을 통하여 외부의 방송 관리 서버로부터 방송 신호 및/또는 방송 관련된 정보를 수신한다.
- [0042] 상기 방송 채널은 위성 채널, 지상파 채널을 포함할 수 있다. 상기 방송 관리 서버는, 방송 신호 및/또는 방송 관련 정보를 생성하여 송신하는 서버 또는 기 생성된 방송 신호 및/또는 방송 관련 정보를 제공받아 단말기에 송신하는 서버를 의미할 수 있다. 상기 방송 신호는, TV 방송 신호, 라디오 방송 신호, 데이터 방송 신호를 포함할 뿐만 아니라, TV 방송 신호 또는 라디오 방송 신호에 데이터 방송 신호가 결합한 형태의 방송 신호도 포함할 수 있다.
- [0043] 상기 방송 관련 정보는, 방송 채널, 방송 프로그램 또는 방송 서비스에 관련한 정보를 의미할 수 있다. 상기 방송 관련 정보는, 이동통신망을 통하여도 제공될 수 있다. 이러한 경우에는 상기 이동통신 모듈(1112)에 의해 수신될 수 있다.
- [0044] 상기 방송 관련 정보는 다양한 형태로 존재할 수 있다. 예를 들어, DMB(Digital Multimedia Broadcasting)의 EPG(Electronic Program Guide) 또는 DVB-H(Digital Video Broadcast-Handheld)의 ESG(Electronic Service Guide) 등의 형태로 존재할 수 있다.
- [0045] 상기 방송 수신 모듈(1111)은, 예를 들어, DMB-T(Digital Multimedia Broadcasting-Terrestrial), DMB-S(Digital Multimedia Broadcasting-Satellite), MediaFLO(Media Forward Link Only), DVB-H(Digital Video Broadcast-Handheld), DVB-CBMS, OMA-BCAST, ISDB-T(Integrated Services Digital Broadcast-Terrestrial) 등의 디지털 방송 시스템을 이용하여 디지털 방송 신호를 수신할 수 있다. 물론, 상기 방송 수신 모듈(1111)은, 상술한 디지털 방송 시스템뿐만 아니라 다른 방송 시스템에 적합하도록 구성될 수도 있다.
- [0046] 방송 수신 모듈(1111)을 통해 수신된 방송 신호 및/또는 방송 관련 정보는 메모리(1160)에 저장될 수 있다.
- [0047] 이동통신 모듈(1112)은, 이동 통신망 상에서 기지국, 외부의 단말, 서버 중 적어도 하나와 무선 신호를 송수신한다. 상기 무선 신호는, 음성 호 신호, 화상 통화 호 신호 또는 문자/멀티미디어 메시지 송수신에 따른 다양한

형태의 데이터를 포함할 수 있다.

- [0048] 무선 인터넷 모듈(1113)은 무선 인터넷 접속을 위한 모듈을 말하는 것으로, 단말기(1100)에 내장되거나 외장될 수 있다.
- [0049] 상기 무선 인터넷의 기술로는 WLAN(Wireless LAN)(Wi-Fi), Wibro(Wireless broadband), Wimax(World Interoperability for Microwave Access), HSDPA(High Speed Downlink Packet Access) 등이 이용될 수 있다.
- [0050] 근거리 통신 모듈(1114)은 근거리 통신을 위한 모듈을 말한다. 상기 근거리 통신(short range communication)의 기술로는 블루투스(Bluetooth), RFID(Radio Frequency Identification), 적외선 통신(IrDA, infrared Data Association), UWB(Ultra Wideband), ZigBee 등이 이용될 수 있다.
- [0051] 위치 정보 모듈(1115)은 단말기(1100)의 위치를 획득하기 위한 모듈로서, 그의 대표적인 예로는 GPS(Global Position System) 모듈이 있다. 현재 기술에 의하면, 상기 GPS모듈(1115)은 3개 이상의 위성으로부터 떨어진 거리 정보와 정확한 시간 정보를 산출한 다음 상기 산출된 정보에 삼각법을 적용함으로써, 위도, 경도, 및 고도에 따른 3차원의 현 위치 정보를 정확히 산출할 수 있다. 현재, 3개의 위성을 이용하여 위치 및 시간 정보를 산출하고, 또 다른 1개의 위성을 이용하여 상기 산출된 위치 및 시간 정보의 오차를 수정하는 방법이 널리 사용되고 있다. 또한, GPS 모듈(1115)은 현 위치를 실시간으로 계속 산출함으로써 속도 정보를 산출할 수 있다.
- [0052] 도 1를 참조하면, A/V(Audio/Video) 입력부(1120)는 오디오 신호 또는 비디오 신호 입력을 위한 것으로, 이에 카메라 모듈(1121)과 마이크(1122) 등이 포함될 수 있다. 카메라 모듈(1121)은 화상 통화모드 또는 촬영 모드에서 이미지 센서에 의해 얻어지는 정지영상 또는 동영상 등의 화상 프레임을 처리한다. 처리된 화상 프레임은 디스플레이부(1151)에 표시될 수 있다.
- [0053] 카메라 모듈(1121)에서 처리된 화상 프레임은 메모리(1160)에 저장되거나 무선 통신부(1110)를 통하여 외부로 전송될 수 있다.
- [0054] 카메라 모듈(1121)은 사용 환경에 따라 2개 이상이 구비될 수도 있다.
- [0055] 일례로, 상기 카메라 모듈(1121)은 단말기(1100)의 디스플레이부(1151)가 구비된 반대면에 3D 영상 촬영을 위한 제1 및 제2 카메라(1121a, 1121b)가 구비될 수 있고, 상기 단말기(1100)의 디스플레이부(1151)가 구비된 면의 일부 영역에 사용자의 셀프 촬영을 위한 제3 카메라(1121c)가 구비될 수 있다.
- [0056] 이때, 제1 카메라(1121a)는 3D 영상의 소스 영상인 좌안 영상 촬영을 위한 것이고, 제2 카메라(1121b)는 우안 영상 촬영을 위한 것이 될 수 있다.
- [0057] 마이크(1122)는 통화모드 또는 녹음모드, 음성 인식 모드 등에서 마이크로폰(Microphone)에 의해 외부의 음향 신호를 입력받아 전기적인 음성 데이터로 처리한다. 처리된 음성 데이터는 통화 모드인 경우 이동통신 모듈(1112)을 통하여 이동통신 기지국으로 송신 가능한 형태로 변환되어 출력될 수 있다. 마이크(1122)에는 외부의 음향 신호를 입력받는 과정에서 발생하는 잡음(noise)을 제거하기 위한 다양한 잡음 제거 알고리즘이 구현될 수 있다.
- [0058] 특히, 마이크(1122)는 상기 음성 인식 모드에서 본 발명의 단말기를 사용하는 사용자의 음성을 녹음하기 위하여 사용될 수 있다. 마이크(1122)에 의하여 녹음된 음성은 제어부(1180)의 감정 인식 수단에 의하여 사용자의 감정 인식에 사용될 수 있다.
- [0059] 입력부(1130)는 사용자가 단말기의 동작 제어를 위한 입력 데이터를 발생시킬 수 있으며, 본 발명의 단말기(1100)에 내장 설치되어 있는 감정 인식 어플리케이션의 전반적인 동작과 관련된 명령을 발생시킬 수 있다.
- [0060] 예를 들어, 사용자가 입력부(1130)를 이용하여 감정 인식 어플리케이션의 실행 명령을 입력하는 경우, 감정 인식 수단의 감정 인식을 위한 일련의 프로세스의 진행을 위하여 디스플레이부(1151)에 소정의 실행창을 표시시킬 수 있다.
- [0061] 입력부(1130)는 본 발명에 따라 표시되는 콘텐츠들 중 두 개 이상의 콘텐츠를 지정하는 신호를 사용자로부터 수신할 수 있다. 그리고, 두 개 이상의 콘텐츠를 지정하는 신호는, 터치입력을 통하여 수신되거나, 하드키 및 소프트 키입력을 통하여 수신될 수 있다.
- [0062] 입력부(1130)는 상기 하나 또는 둘 이상의 콘텐츠들을 선택하는 입력을 사용자로부터 수신할 수 있다. 또한, 사용자로부터 단말기(1100)가 수행할 수 있는 기능과 관련된 아이콘을 생성하는 입력을 수신할 수 있다.

- [0063] 상기와 같은, 입력부(1130)는 방향키, 키 패드(key pad), 돔 스위치 (dome switch), 터치 패드(정압/정전), 조그 휠, 조그 스위치 등으로 구성될 수 있다.
- [0064] 센싱부(1140)는 단말기(1100)의 개폐 상태, 단말기(1100)의 위치, 사용자 접촉 유무, 단말기(1100)의 방위, 단말기(1100)의 가속/감속 등과 같이 단말기(1100)의 현 상태를 감지하여 단말기(1100)의 동작을 제어하기 위한 센싱 신호를 발생시킨다. 예를 들어 단말기(1100)가 슬라이드 폰 형태인 경우 슬라이드 폰의 개폐 여부를 센싱할 수 있다. 또한, 배터리의 전원 공급 여부, 인터페이스부(1170)의 외부 기기 결합 여부 등을 센싱할 수도 있다. 한편, 상기 센싱부(1140)는 근접 센서(1141)를 포함할 수 있다. 상기 근접 센서(1141)에 대해서는 나중에 터치스크린과 관련되어 후술된다.
- [0065] 출력부(1150)는 시각, 청각 또는 촉각 등과 관련된 출력을 발생시키기 위한 것으로, 이에는 디스플레이부(1151), 음향 출력 모듈(1152), 알람부(1153), 햅틱 모듈(1154) 및 프로젝터 모듈(1155) 등이 포함될 수 있다.
- [0066] 디스플레이부(1151)는 단말기(1100)에서 처리되는 정보를 표시(출력)한다. 예를 들어, 단말기가 통화 모드인 경우 통화와 관련된 UI(User Interface) 또는 GUI(Graphic User Interface)를 표시한다. 단말기(1100)가 화상 통화 모드 또는 촬영 모드인 경우에는 촬영 또는/및 수신된 영상 또는 UI, GUI를 표시한다.
- [0067] 또한, 본 발명에 따른 디스플레이부(1151)는 2D 및 3D 표시 모드를 지원한다.
- [0068] 즉, 본 발명에 따른 디스플레이부(1151)는 이하의 도 1에 도시된 바와 같이, 일반적인 디스플레이 장치(1151a)에 스위치 액정(1151b)을 조합하는 구성을 가질 수 있다. 그리고, 스위치 액정(1151b)을 이용하여 광학 시차 장벽을 작동시켜 광의 진행 방향을 제어하여 좌우의 눈에 각기 다른 광이 도달하도록 분리할 수 있다. 때문에 우안용 영상과 좌안용 영상이 조합된 영상이 디스플레이 장치(1151a)에 표시되는 경우 사용자의 입장에서는 각각의 눈에 대응한 화상이 보여 마치 입체로 표시된 것처럼 느끼게 된다.
- [0069] 즉, 디스플레이부(1151)는 제어부(1180)의 제어에 따라, 2D 표시 모드인 상태에서는 상기 스위치 액정(1151b) 및 광학 시차 장벽을 구동시키지 않고, 상기 디스플레이 장치(1151a)만을 구동시켜 일반적인 2D 표시 동작을 수행한다.
- [0070] 또한, 디스플레이부(1151)는 제어부(1180)의 제어에 따라, 3D 표시 모드인 상태에서는 상기 스위치 액정(1151b)과, 광학 시차 장벽 및 디스플레이 장치(1151a)를 구동시켜 3D 표시 동작을 수행한다.
- [0071] 한편, 상기와 같은 디스플레이부(1151)는 액정 디스플레이(liquid crystal display, LCD), 박막 트랜지스터 액정 디스플레이(thin film transistor-liquid crystal display, TFT LCD), 유기 발광 다이오드(organic light-emitting diode, OLED), 플렉시블 디스플레이(flexible display), 3차원 디스플레이(3D display) 중에서 적어도 하나를 포함할 수 있다.
- [0072] 이들 중 일부 디스플레이는 그를 통해 외부를 볼 수 있도록 투명형 또는 광투과형으로 구성될 수 있다. 이는 투명 디스플레이라 호칭될 수 있는데, 상기 투명 디스플레이의 대표적인 예로는 TOLED(Transparent OLED) 등이 있다. 디스플레이부(1151)의 후방 구조 또한 광 투과형 구조로 구성될 수 있다. 이러한 구조에 의하여, 사용자는 단말기 바디의 디스플레이부(1151)가 차지하는 영역을 통해 단말기 바디의 후방에 위치한 사물을 볼 수 있다.
- [0073] 단말기(1100)의 구형 형태에 따라 디스플레이부(1151)가 2개 이상 존재할 수 있다. 예를 들어, 단말기(1100)에는 복수의 디스플레이부들이 하나의 면에 이격되거나 일체로 배치될 수 있고, 또한 서로 다른 면에 각각 배치될 수도 있다.
- [0074] 디스플레이부(1151)와 터치 동작을 감지하는 센서(이하, '터치 센서'라 함)가 상호 레이어 구조를 이루는 경우(이하, '터치 스크린'이라 함)에, 디스플레이부(1151)는 출력 장치 이외에 입력 장치로도 사용될 수 있다. 터치 센서는, 예를 들어, 터치 필름, 터치 시트, 터치 패드 등의 형태를 가질 수 있다.
- [0075] 터치 센서는 디스플레이부(1151)의 특정 부위에 가해진 압력 또는 디스플레이부(1151)의 특정 부위에 발생하는 정전 용량 등의 변화를 전기적인 입력신호로 변환하도록 구성될 수 있다. 터치 센서는 터치 되는 위치 및 면적 뿐만 아니라, 터치 시의 압력까지도 검출할 수 있도록 구성될 수 있다.
- [0076] 터치 센서에 대한 터치 입력이 있는 경우, 그에 대응하는 신호(들)는 터치 제어기(미도시)로 보내진다. 터치 제어기는 그 신호(들)를 처리한 다음 대응하는 데이터를 제어부(1180)로 전송한다. 이로써, 제어부(1180)는 디스플레이부(1151)의 어느 영역이 터치 되었는지 여부 등을 알 수 있게 된다.
- [0077] 상기 근접 센서(1141)는 상기 터치스크린에 의해 감싸지는 단말기(1100)의 내부 영역 또는 상기 터치 스크린의

근처에 배치될 수 있다. 상기 근접 센서는 소정의 검출면에 접근하는 물체, 혹은 근방에 존재하는 물체의 유무를 전자계의 힘 또는 적외선을 이용하여 기계적 접촉이 없이 검출하는 센서를 말한다. 근접 센서는 접촉식 센서 보다는 그 수명이 길며 그 활용도 또한 높다.

[0078] 상기 근접 센서의 예로는 투과형 광전 센서, 직접 반사형 광전 센서, 미러 반사형 광전 센서, 고주파 발진형 근접 센서, 정전용량형 근접 센서, 자기형 근접 센서, 적외선 근접 센서 등이 있다. 상기 터치스크린이 정전식인 경우에는 상기 포인터의 근접에 따른 전계의 변화로 상기 포인터의 근접을 검출하도록 구성된다. 이 경우 상기 터치 스크린(터치 센서)은 근접 센서로 분류될 수도 있다.

[0079] 이하에서는 설명의 편의를 위해, 상기 터치스크린 상에 포인터가 접촉되지 않으면서 근접되어 상기 포인터가 상기 터치스크린 상에 위치함이 인식되도록 하는 행위를 "근접 터치(proximity touch)"라고 칭하고, 상기 터치스크린 상에 포인터가 실제로 접촉되는 행위를 "접촉 터치(contact touch)"라고 칭한다. 상기 터치스크린 상에서 포인터로 근접 터치가 되는 위치라 함은, 상기 포인터가 근접 터치될 때 상기 포인터가 상기 터치스크린에 대해 수직으로 대응되는 위치를 의미한다.

[0080] 상기 근접센서는, 근접 터치와, 근접 터치 패턴(예를 들어, 근접 터치 거리, 근접 터치 방향, 근접 터치 속도, 근접 터치 시간, 근접 터치 위치, 근접 터치 이동 상태 등)을 감지한다. 상기 감지된 근접 터치 동작 및 근접 터치 패턴에 상응하는 정보는 터치 스크린상에 출력될 수 있다.

[0081] 음향 출력 모듈(1152)은 호신호 수신, 통화모드 또는 녹음 모드, 음성인식 모드, 방송수신 모드 등에서 무선 통신부(1110)로부터 수신되거나 메모리(1160)에 저장된 오디오 데이터를 출력할 수 있다. 음향 출력 모듈(1152)은 단말기(1100)에서 수행되는 기능(예를 들어, 호신호 수신음, 메시지 수신음 등)과 관련된 음향 신호를 출력하기도 한다. 이러한 음향 출력 모듈(1152)에는 리시버(Receiver), 스피커(speaker), 버저(Buzzer) 등이 포함될 수 있다.

[0082] 알람부(1153)는 단말기(1100)의 이벤트 발생을 알리기 위한 신호를 출력한다. 단말기(1100)에서 발생 되는 이벤트의 예로는 호 신호 수신, 메시지 수신, 키 신호 입력, 터치 입력 등이 있다. 알람부(1153)는 비디오 신호나 오디오 신호 이외에 다른 형태, 예를 들어 진동으로 이벤트 발생을 알리기 위한 신호를 출력할 수도 있다. 상기 비디오 신호나 오디오 신호는 디스플레이부(1151)나 음성 출력 모듈(1152)을 통해서도 출력될 수 있으므로, 이 경우 상기 디스플레이부(1151) 및 음향 출력 모듈(1152)은 알람부(1153)의 일종으로 분류될 수도 있다.

[0083] 햅틱 모듈(haptic module)(1154)은 사용자가 느낄 수 있는 다양한 촉각 효과를 발생시킨다. 햅틱 모듈(1154)이 발생시키는 촉각 효과의 대표적인 예로는 진동이 있다. 햅틱 모듈(1154)이 발생하는 진동의 세기와 패턴 등은 제어가능하다. 예를 들어, 서로 다른 진동을 합성하여 출력하거나 순차적으로 출력할 수도 있다.

[0084] 햅틱 모듈(1154)은, 진동 외에도, 접촉 피부면에 대해 수직 운동하는 핀 배열, 분사구나 흡입구를 통한 공기의 분사력이나 흡입력, 피부 표면에 대한 스침, 전극(electrode)의 접촉, 정전기력 등의 자극에 의한 효과와, 흡열이나 발열 가능한 소자를 이용한 냉온감 재현에 의한 효과 등 다양한 촉각 효과를 발생시킬 수 있다.

[0085] 햅틱 모듈(1154)은 직접적인 접촉을 통해 촉각 효과의 전달할 수 있을 뿐만 아니라, 사용자가 손가락이나 팔 등의 근 감각을 통해 촉각 효과를 느낄 수 있도록 구현할 수도 있다. 햅틱 모듈(1154)은 단말기(1100)의 구성 태양에 따라 2개 이상이 구비될 수 있다.

[0086] 프로젝터 모듈(1155)은, 단말기(1100)를 이용하여 이미지 프로젝트(project) 기능을 수행하기 위한 구성요소로서, 제어부(1180)의 제어 신호에 따라 디스플레이부(1151)상에 디스플레이되는 영상과 동일하거나 적어도 일부가 다른 영상을 외부 스크린 또는 벽에 디스플레이할 수 있다.

[0087] 구체적으로, 프로젝터 모듈(1155)은, 영상을 외부로 출력하기 위한 빛(일 예로서, 레이저 광)을 발생시키는 광원(미도시), 광원에 의해 발생한 빛을 이용하여 외부로 출력할 영상을 생성하기 위한 영상 생성 수단(미도시), 및 영상을 일정 초점 거리에서 외부로 확대 출력하기 위한 렌즈(미도시)를 포함할 수 있다. 또한, 프로젝터 모듈(1155)은, 렌즈 또는 모듈 전체를 기계적으로 움직여 영상 투사 방향을 조절할 수 있는 장치(미도시)를 포함할 수 있다.

[0088] 프로젝터 모듈(1155)은 디스플레이 수단의 소자 종류에 따라 CRT(Cathode Ray Tube) 모듈, LCD(Liquid Crystal Display) 모듈 및 DLP(Digital Light Processing) 모듈 등으로 나뉠 수 있다. 특히, DLP 모듈은, 광원에서 발생한 빛이 DMD(Digital Micromirror Device) 칩에 반사됨으로써 생성된 영상을 확대 투사하는 방식으로 프로젝터 모듈(1155)의 소형화에 유리할 수 있다.

- [0089] 바람직하게, 프로젝터 모듈(1155)은, 단말기(1100)의 측면, 정면 또는 배면에 길이 방향으로 구비될 수 있다. 물론, 프로젝터 모듈(1155)은, 필요에 따라 단말기(1100)의 어느 위치에라도 구비될 수 있음은 당연하다.
- [0090] 또한, 출력부(1150)는 후각정보를 발생시킬 수 있는 방향부를 더 포함할 수 있다. 방향부는 단말기(1100) 외부로 소정의 향기나 냄새를 발생시킬 수 있다.
- [0091] 메모리(1160)는 제어부(1180)의 처리 및 제어를 위한 프로그램이 저장될 수도 있고, 입/출력되는 데이터들(예를 들어, 전화번호부, 메시지, 오디오, 정지영상, 전자도서, 동영상, 송수신 메시지 히스토리 등)의 임시 저장을 위한 기능을 수행할 수도 있다. 상기 메모리(1160)에는 상기 데이터들 각각에 대한 사용 빈도(예를 들면, 각 전화번호, 각 메시지, 각 멀티미디어에 대한 사용빈도)도 함께 저장될 수 있다. 또한, 상기 메모리(1160)에는 상기 터치스크린 상의 터치 입력시 출력되는 다양한 패턴의 진동 및 음향에 관한 데이터를 저장할 수 있다.
- [0092] 또한, 메모리(1160)는 복수의 감정에 대한 각 감정별 음성 데이터로부터 훈련된 은닉 마르코프 모델(Hidden Markov Model)과 같은 음향 모델 데이터를 저장할 수 있다. 이 같은 음향 모델 데이터는 감정 인식 수단에 의하여 사용자의 감정을 판단하는 데 사용된다.
- [0093] 또한, 메모리(1160)에는 사용자의 음성에 대한 특징 파라미터를 추가적으로 저장함으로써 음향 모델 데이터를 업데이트하는 데 사용할 수 있다. 특징 파라미터는 멜 주파수 켈프스트럼 계수(Mel-Frequency Cepstral Coefficient) 등의 음향 파라미터를 나타낸다.
- [0094] 상기와 같은 메모리(1160)는 플래시 메모리 타입(flash memory type), 하드디스크 타입(hard disk type), 멀티미디어 카드 마이크로 타입(multimedia card micro type), 카드 타입의 메모리(예를 들어 SD 또는 XD 메모리 등), 램(Random Access Memory, RAM), SRAM(Static Random Access Memory), 롬(Read-Only Memory, ROM), EEPROM(Electrically Erasable Programmable Read-Only Memory), PROM(Programmable Read-Only Memory), 자기 메모리, 자기 디스크, 광디스크 중 적어도 하나의 타입의 저장매체를 포함할 수 있다. 단말기(1100)는 인터넷(internet)상에서 상기 메모리(1160)의 저장 기능을 수행하는 웹 스토리지(web storage)와 관련되어 동작할 수도 있다.
- [0095] 인터페이스부(1170)는 단말기(1100)에 연결되는 모든 외부기기와의 통로 역할을 한다. 인터페이스부(1170)는 외부 기기로부터 데이터를 전송받거나, 전원을 공급받아 단말기(1100) 내부의 각 구성 요소에 전달하거나, 단말기(1100) 내부의 데이터가 외부 기기로 전송되도록 한다. 예를 들어, 유/무선 헤드셋 포트, 외부 충전기 포트, 유/무선 데이터 포트, 메모리 카드(memory card) 포트, 식별 모듈이 구비된 장치를 연결하는 포트, 오디오 I/O(Input/Output) 포트, 비디오 I/O(Input/Output) 포트, 이어폰 포트 등이 인터페이스부(1170)에 포함될 수 있다.
- [0096] 식별 모듈은 단말기(1100)의 사용 권한을 인증하기 위한 각종 정보를 저장한 칩으로서, 사용자 인증 모듈(User Identify Module, UIM), 가입자 인증 모듈(Subscriber Identify Module, SIM), 범용 사용자 인증 모듈(Universal Subscriber Identity Module, USIM) 등을 포함할 수 있다. 식별 모듈이 구비된 장치(이하 '식별 장치')는, 스마트 카드(smart card) 형식으로 제작될 수 있다. 따라서 식별 장치는 포트를 통하여 단말기(1100)와 연결될 수 있다.
- [0097] 상기 인터페이스부(1170)는 단말기(1100)가 외부 크래들(cradle)과 연결될 때 상기 크래들로부터의 전원이 상기 단말기(1100)에 공급되는 통로가 되거나, 사용자에게 의해 상기 크래들에서 입력되는 각종 명령 신호가 상기 단말기(1100)로 전달되는 통로가 될 수 있다. 상기 크래들로부터 입력되는 각종 명령 신호 또는 상기 전원은 상기 단말기(1100)가 상기 크래들에 정확히 장착되었음을 인지하기 위한 신호로 동작될 수도 있다.
- [0098] 제어부(controller)(1180)는 통상적으로 단말기(1100)의 전반적인 동작을 제어한다. 예를 들어 음성 통화, 데이터 통신, 화상 통화 등을 위한 관련된 제어 및 처리를 수행한다. 제어부(1180)는 멀티 미디어 재생을 위한 멀티 미디어 모듈(1181)을 구비할 수도 있다. 멀티미디어 모듈(1181)은 제어부(1180) 내에 구현될 수도 있고, 제어부(1180)와 별도로 구현될 수도 있다.
- [0099] 상기 제어부(1180)는 상기 터치스크린 상에서 행해지는 필기 입력 또는 그림 그리기 입력을 각각 문자 및 이미지로 인식할 수 있는 패턴 인식 처리를 행할 수 있다.
- [0100] 특히, 제어부(1180)는 본 발명의 단말기(1100)가 사용자의 음성을 이용하여 감정을 인식하기 위한 감정 인식 수단을 포함할 수 있다. 상기 감정 인식 수단은 신호 에너지(signal energy) 기반의 voice activity detection으로서 사용자의 음성을 감지하는 감지부, 특징 파라미터를 추출하는 추출부 및 사용자의 감정을 판단하는 인식부

등으로 구성될 수 있다.

- [0101] 한편, 제어부(1180)는 상기 디스플레이부(1151)가 유기 발광 다이오드(organic light-emitting diode, OLED) 또는 TOLED(Transparent OLED)로 구비될 경우, 본 발명에 따라, 카메라(1121)를 통해 입력된 프리뷰 영상이 상기 유기 발광 다이오드(organic light-emitting diode, OLED) 또는 TOLED(Transparent OLED)의 화면 상에 풀업 표시된 상태에서, 사용자에게 조작에 따라 상기 프리뷰 영상의 크기가 조절되면, 상기 화면 상에서 상기 크기가 조절된 프리뷰 영상이 표시된 제1 영역을 제외한 나머지 제2 영역 내의 화소들의 구동을 오프시킴으로써, 전원 공급부(1190)에서 상기 디스플레이부(1151)로 공급되는 전원의 소모량을 줄일 수 있다.
- [0102] 전원 공급부(1190)는 제어부(1180)의 제어에 의해 외부의 전원, 내부의 전원을 인가받아 각 구성요소들의 동작에 필요한 전원을 공급한다.
- [0103] 여기에 설명되는 다양한 실시례는 예를 들어, 소프트웨어, 하드웨어 또는 이들의 조합된 것을 이용하여 컴퓨터 또는 이와 유사한 장치로 읽을 수 있는 기록매체 내에서 구현될 수 있다.
- [0104] 하드웨어적인 구현에 의하면, 여기에 설명되는 실시례는 ASICs (application specific integrated circuits), DSPs (digital signal processors), DSPDs (digital signal processing devices), PLDs (programmable logic devices), FPGAs (field programmable gate arrays, 프로세서(processors), 제어기(controllers), 마이크로 컨트롤러(micro-controllers), 마이크로 프로세서(microprocessors), 기타 기능 수행을 위한 전기적인 유닛 중 적어도 하나를 이용하여 구현될 수 있다. 일부의 경우에 본 명세서에서 설명되는 실시례들이 제어부(1180) 자체로 구현될 수 있다.
- [0105] 소프트웨어적인 구현에 의하면, 본 명세서에서 설명되는 절차 및 기능과 같은 실시예들은 별도의 소프트웨어 모듈들로 구현될 수 있다. 상기 소프트웨어 모듈들 각각은 본 명세서에서 설명되는 하나 이상의 기능 및 작동을 수행할 수 있다. 적절한 프로그램 언어로 쓰여진 소프트웨어 어플리케이션으로 소프트웨어 코드가 구현될 수 있다. 상기 소프트웨어 코드는 메모리(1160)에 저장되고, 제어부(1180)에 의해 실행될 수 있다.
- [0106] 도 2a 및 도 2b는 본 발명과 관련된 단말기의 전면 사시도 및 후면 사시도의 일례를 나타낸다.
- [0107] 도 2a를 참조하면, 개시된 단말기(1100)는 바 형태의 본체를 구비하고 있다. 다만, 본 발명은 여기에 한정되지 않고, 2 이상의 바들이 상대 이동 가능하게 결합되는 슬라이드 타입, 폴더 타입, 스윙 타입, 스위블 타입 등 다양한 구조에 적용이 가능하다.
- [0108] 바디는 외관을 이루는 케이스(케이싱, 하우징, 커버 등)를 포함한다. 본 실시예에서, 케이스는 프론트 케이스(1101)와 리어 케이스(1102)로 구분될 수 있다. 프론트 케이스(1101)와 리어 케이스(1102)의 사이에 형성된 공간에는 각종 전자부품들이 내장된다. 프론트 케이스(1101)와 리어 케이스(1102) 사이에는 적어도 하나의 중간 케이스가 추가로 배치될 수도 있다.
- [0109] 케이스들은 합성수지를 사출하여 형성되거나 금속 재질, 예를 들어 스테인레스 스틸(STS) 또는 티타늄(Ti) 등과 같은 금속 재질을 갖도록 형성될 수도 있다.
- [0110] 단말기 바디, 주로 프론트 케이스(1101)에는 디스플레이부(1151), 음향 출력 모듈(1152), 제3 카메라(1121c), 입력부(1130, 1131, 1132), 마이크(1122), 인터페이스부(1170) 등이 배치될 수 있다.
- [0111] 디스플레이부(1151)는 프론트 케이스(1101)의 주면의 대부분을 차지한다. 디스플레이부(1151)의 양단부 중 일단부에 인접한 영역에는 음향 출력 모듈(1151)과 카메라 모듈(1121)이 배치되고, 다른 단부에 인접한 영역에는 입력부(1131)와 마이크(1122)가 배치된다. 입력부(1132)와 인터페이스부(1170) 등은 프론트 케이스(1101) 및 리어 케이스(1102)의 측면들에 배치될 수 있다.
- [0112] 제3 카메라(1121c)는 제1 및 제2 카메라(1121a, 1121b)와 실질적으로 반대되는 촬영 방향을 가지며, 제1 및 제2 카메라(1121a, 1121b)와 같거나 서로 다른 화소를 가지는 카메라일 수 있다.
- [0113] 제3 카메라(1121c)에 인접하게는 플래쉬와 거울이 추가로 배치될 수 있다. 플래쉬는 제3 카메라(1121c)로 피사체를 촬영하는 경우에 피사체를 향해 빛을 비추게 된다. 거울은 사용자가 제3 카메라(1121c)를 이용하여 자신을 촬영(셀프 촬영)하고자 하는 경우에, 사용자 자신의 얼굴 등을 비춰볼 수 있게 한다.
- [0114] 입력부(1130)는 단말기(1100)의 동작을 제어하기 위한 명령을 입력받기 위해 조작되는 것으로서, 복수의 조작 유닛들(1131, 1132)을 포함할 수 있다. 조작 유닛들(1131, 1132)은 조작부(manipulating portion)로도 통칭될 수 있다.

- [0115] 제1 또는 제2 조작 유닛들(1131, 1132)에 의하여 입력되는 내용은 다양하게 설정될 수 있다. 예를 들어, 제1 조작 유닛(1131)은 시작, 종료, 스크롤 등과 같은 명령을 입력받고, 제2 조작 유닛(1132)은 음향 출력 모듈(1152)에서 출력되는 음향의 크기 조절 또는 디스플레이부(1151)의 터치 인식 모드 활성화/비활성화 등과 같은 명령을 입력받을 수 있다.
- [0116] 마이크(1122)는 외부의 음향 신호를 입력받아 전기적인 음성 데이터로 처리는 구성으로서, 특히 사용자의 음성을 녹음하기 위하여 사용될 수 있으며, 이는 본 발명에 적용될 수 있는 사용자의 감정 인식에 이용될 수 있다.
- [0117] 도 2b를 참조하면, 단말기 바디의 후면, 다시 말해서 리어 케이스(1102)에는 제3 카메라(1121c)와 반대 방향을 촬영하는 제1 및 제2 카메라(1121a, 1121b)가 추가로 장착될 수 있다.
- [0118] 단말기 바디의 후면에는 음향 출력 모듈(1152')이 추가로 배치될 수도 있다. 음향 출력부(1152')는 음향 출력 모듈(1152, 도 2a 참조)과 함께 스테레오 기능을 구현할 수 있으며, 통화시 스피커폰 모드의 구현을 위하여 사용될 수도 있다.
- [0119] 통화 등을 위한 안테나 외에 방송신호 수신용 안테나(1116)가 휴대 단말기 바디의 측면에 추가적으로 배치될 수 있다. 방송 수신부(1111, 도 1 참조)의 일부를 이루는 안테나(1116)는 단말기 바디에서 인출 가능하게 설치될 수 있다.
- [0120] 단말기 바디에는 단말기(1100)에 전원을 공급하기 위한 전원 공급부(1190)가 장착된다. 전원 공급부(1190)는 단말기 바디에 내장되거나, 단말기 바디의 외부에서 직접 탈착될 수 있게 구성될 수 있다.
- [0121] 리어 케이스(1102)에는 터치를 감지하기 위한 터치 패드가 추가로 장착될 수 있다. 터치 패드는 디스플레이부(1151)를 위해 광 투과형으로 구성될 수도 있다. 이 경우에, 디스플레이부가 양면에서(즉, 이동단말기의 전면 및 후면 둘 다의 방향으로) 시각 정보를 출력하도록 구성된다면, 터치 패드를 통해서도 상기 시각 정보를 인지할 수 있게 된다. 상기 양면에 출력되는 정보는 상기 터치 패드에 의해 모두 제어될 수도 있다.
- [0122] 한편, 터치 패드 전용 디스플레이가 별도로 장착됨으로써, 리어 케이스(1102)에도 터치 스크린이 배치될 수도 있다.
- [0123] 터치 패드는 프론트 케이스(1101)의 디스플레이부(1151)와 상호 관련되어 작동한다. 터치 패드는 디스플레이부(1151)의 후방에 평행하게 배치될 수 있다. 이러한 터치 패드는 디스플레이부(1151)와 동일하거나 작은 크기를 가질 수 있다.
- [0124] 이하, 도 3을 참조하여 디스플레이부(1151)와 터치 패드의 서로 연관된 작동 방식에 대하여 살펴본다. 도 3은 본 발명과 관련된 단말기의 일 작동 상태를 설명하기 위한 단말기의 정면도이다.
- [0125] 디스플레이부(1151)에는 다양한 종류의 시각 정보들이 표시될 수 있다. 이들 정보들은 문자, 숫자, 기호, 그래픽, 또는 아이콘 등의 형태로 표시될 수 있다.
- [0126] 이러한 정보의 입력을 위하여 상기 문자, 숫자, 기호, 그래픽 또는 아이콘 들 중 적어도 하나는 일정한 배열을 이루어 표시됨으로써 키패드의 형태로 구현될 수 있다. 이러한 키패드는 소위 '가상 키패드'(virtual keypad)라 불릴 수 있다.
- [0127] 도 3은 단말기 바디의 전면을 통해 가상 키패드에 가해진 터치를 입력받는 것을 나타내고 있다.
- [0128] 디스플레이부(1151)는 전체 영역으로 작동되거나, 복수의 영역들로 나뉘어져 작동될 수 있다. 후자의 경우, 상기 복수의 영역들은 서로 연관되게 작동되도록 구성될 수 있다.
- [0129] 예를 들어, 디스플레이부(1151)의 상부와 하부에는 출력창(1151a)과 입력창(1151b)이 각각 표시된다. 출력창(1151a)과 입력창(1151b)은 각각 정보의 출력 또는 입력을 위해 할당되는 영역이다. 입력창(1151b)에는 전화 번호 등의 입력을 위한 숫자가 표시된 가상 키패드(1151c)가 출력된다. 가상 키패드(1151c)가 터치되면, 터치된 가상 키패드에 대응되는 숫자 등이 출력창(1151a)에 표시된다. 제1조작 유닛(1131)이 조작되면 출력창(1151a)에 표시된 전화번호에 대한 호 연결이 시도된다.
- [0130] 이상의 실시예들에 개시된 입력 방식뿐만 아니라, 디스플레이부(1151) 또는 터치 패드는 스크롤(scroll)에 의해 터치 입력받도록 구성될 수 있다. 사용자는 디스플레이부(1151) 또는 터치 패드를 스크롤 함으로써 디스플레이부(1151)에 표시된 객체, 예를 들어 아이콘 등에 위치한 커서 또는 포인터를 이동시킬 수 있다. 나아가, 손가락을 디스플레이부(1151) 또는 터치 패드 상에서 이동시키는 경우, 손가락이 움직이는 경로가 디스플레이부(1151)

1)에 시각적으로 표시될 수도 있다. 이는 디스플레이부(1151)에 표시되는 이미지를 편집함에 유용할 것이다.

- [0131] 디스플레이부(1151)(터치 스크린) 및 터치 패드가 일정 시간 범위 내에서 함께 터치되는 경우에 대응하여, 단말기의 일 기능이 실행될 수도 있다. 함께 터치되는 경우로는, 사용자가 엄지 및 검지를 이용하여 단말기 바디를 잡는(clamping) 경우가 있을 수 있다. 상기 일 기능은, 예를 들어, 디스플레이부(1151) 또는 터치 패드에 대한 활성화 또는 비활성화 등이 있을 수 있다.
- [0132] 한편, 디스플레이부(1151)에서 특정 오브젝트를 가리키거나 메뉴를 선택하기 위한 화살표 또는 손가락 형태의 그래픽은 포인터(pointer) 또는 커서(cursor)로 호칭된다. 그러나, 포인터의 경우 터치 조작 등을 위한 손가락이나 스타일러스 펜 등을 의미하는 것으로 혼용되는 경우가 많다. 따라서 본 명세서에서는 이 둘을 명확히 구분하기 위하여 디스플레이부에 표시되는 그래픽을 커서라 칭하고, 손가락이나 스타일러스 펜과 같이 터치, 근접터치, 제스처(gesture)를 수행할 수 있는 물리적 수단을 포인터라 칭한다.
- [0133] 본 발명에 따른 음성을 이용한 감정 인식 수단은 감정을 인식하고자 하는 객체, 즉 대상 객체의 음성을 감지하여 대상 객체가 느끼는 4가지 감정(예를 들어, 중립, 기쁨, 분노 및 슬픔)을 분류하여 인식할 수 있다. 단, 분류된 상기 4가지의 감정은 본 발명이 적용되기 위한 단순한 일례에 불과하고 다른 기준에 의해 감정은 복수로 분류되는 것도 가능하다.
- [0134] 본 발명에 적용될 수 있는 감정 인식 수단은 제시된 자극에 대응하여 반응하는 객체의 객체정보를 입력받아 객체의 음성을 인식할 수 있다. 이때 객체는 감정을 인식하고자 하는 대상체이다. 대상체는 사람이 될 수 있으며, 특정 동물의 감정을 인식하기 위해서는 감정을 인식하고자 하는 동물이 될 수도 있다.
- [0135] 객체정보는 피측정 대상체의 정보로서 주로 피측정 대상체의 음성을 녹음한 데이터이다.
- [0136] 객체정보의 일실시례로서 마이크를 통해 대상체의 음성을 녹음한 경우 이로부터 생성된 디지털 데이터가 객체정보일 수 있다.
- [0137] 또한, 감정 인식 수단은 대상체의 음성을 이용하여 대상체의 음성에 대한 특징 파라미터를 추출하며, 감정 각각에 대응하는 음성의 특징 파라미터에 대한 정보를 이용하여 대상체의 감정을 판단할 수 있다.
- [0138] 이하에서는, 본 발명에 적용되는 감정 인식 어플리케이션에 대해 구체적으로 서술한다. 도 4는 본 발명에 따른 감정 인식 어플리케이션의 실행에 의한 실행창의 일 실시례를 나타낸다.
- [0139] 단말기(1100)의 사용자가 입력부(1130)를 이용하여 상기 감정 인식 어플리케이션의 실행 명령을 입력하는 경우, 감정 인식 수단과 관련된 소정의 실행창이 디스플레이부(1151)에 표시된다.
- [0140] 도 4를 참조하면, 실행창의 상단에는 "OnlineEmotionRecognition"이라는 어플리케이션의 명칭이 나타난다.
- [0141] "Start Record" 버튼은 사용자의 음성 녹음을 시작하기 위한 버튼이다. "Start Record" 버튼을 누르면 음성 녹음이 시작되고, 묵음(비음성)이 감지되면 그때까지 입력된 음성에 기초하여 인식된 사용자의 감정이 화면상에 출력된다.
- [0142] "Play Sound" 버튼은 현재까지 녹음된 음성을 재생하기 위한 버튼이다. "Play Sound" 버튼을 누르면 녹음된 음성의 목록이 나타나며, 상기 목록 중 사용자가 선택한 음성이 재생된다.
- [0143] "Male", "Female"은 음성을 입력하는 사용자의 성별을 선택하기 위한 수단이다. 본 발명은 남녀에 따른 감정 인식 특성이 다르므로 감정 인식의 성능을 향상시키기 위하여 성별 감정 모델을 사용한다. 즉, "Male"이 선택된 상태에서 녹음된 음성은 남성 감정 모델을 이용하여 감정이 인식되고, "Female"이 선택된 상태에서 녹음된 음성은 여성 감정 모델을 이용하여 감정이 인식된다.
- [0144] 감정 인식 어플리케이션에 의한 실행창에는 상기 버튼 이외에도 음성의 녹음 시작 시점이나 음성 녹음되는 시간을 설정할 수 있는 선택사항이 별도로 구비될 수 있다. 또한, 실행창에는 업데이트 명령을 위한 버튼이 있어 사용자에게 대한 음성 데이터가 메모리(1160)에 추가될 수 있다.
- [0145] 이하에서는, 전술한 감정 인식 수단을 통해 객체의 음성을 이용하여 감정을 인식하는 방법에 대해 구체적으로 서술한다.
- [0146] 상술한 바와 같이, 감정 인식 수단은 감지부, 추출부 및 인식부 등으로 구성될 수 있다.
- [0147] 감지부는 신호 에너지(signal energy) 기반의 voice activity detection을 수행하며, 사용자로부터 녹음된 제

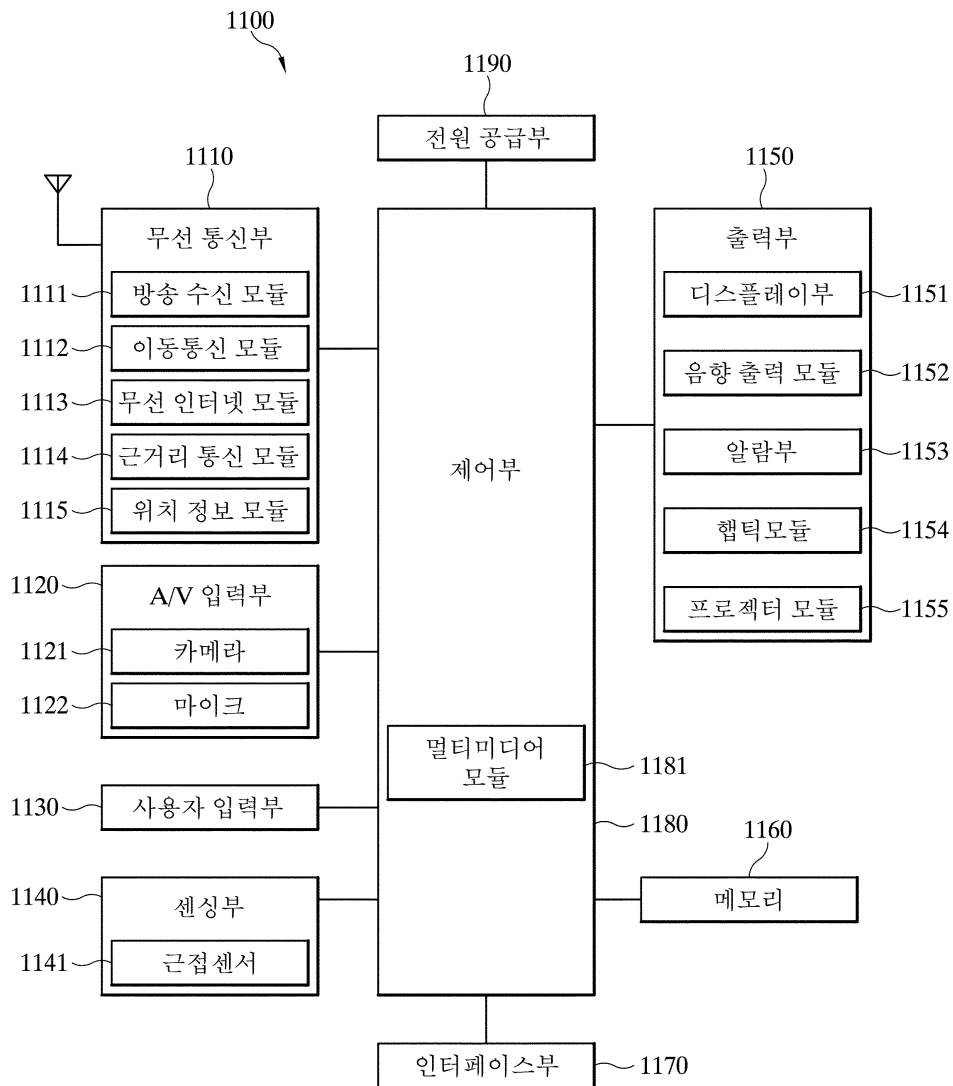
1 음성으로부터 사용자의 음성을 감지할 수 있다.

- [0148] 도 5a 및 도 5b는 감정 인식 수단의 추출부가 특징 파라미터를 추출하는 과정의 일 실시례를 개략적으로 나타낸 것이다.
- [0149] 도 5a 및 도 5b를 참조하면, 추출부는 마이크(1122)에 의하여 녹음된 사용자의 음성이 포함된 음원으로부터 특징 파라미터를 추출하며, 상기 특징 파라미터는 MFCC(Mel-Frequency Cepstral Coefficient)와 Log-energy를 기반으로 한다.
- [0150] 도 5a를 참조하면, MFCC는 Hamming Window를 거친 사용자의 음성에 FFT(Fast Fourier Transform)를 적용하고, 상기 FFT가 적용된 결과에 Mel-scale의 필터 뱅크를 적용하여 파워스펙트럼을 구하며, 상기 파워스펙트럼에 로그(log)를 취하고, 상기 로그를 취한 결과에 DCT(Discrete Cosine Transform)를 적용하여 얻어질 수 있다.
- [0151] 도 5b를 참조하면, 특징 파라미터는 녹음된 음성의 매 프레임마다 39차의 실수열을 추출하여 얻어진다. 여기서 프레임의 길이는 30ms로 설정될 수 있다.
- [0152] 39차의 실수열은 3개의 13차 실수열로 구성된다. 첫번째 13차 실수열은 현재 프레임에서 추출한 MFCC(12차)와 Log-energe(1차)를 이용하여 얻어지고, 두번째 13차 실수열은 현재 프레임과 1차 이전 프레임과의 요소별 차분을 이용하여 얻어지며, 세번째 13차 실수열은 현재 프레임과 2차 이전 프레임과의 요소별 차분을 이용하여 얻어진다.
- [0153] 도 6a 및 도 6b는 감정 인식 수단의 인식부가 감정을 훈련 및 인식하는 과정의 일 실시례를 개략적으로 나타낸 것이고, 도 7은 본 발명의 감정 인식 수단에 따른 인식률 및 인식시간에 대한 실험 결과를 나타낸다.
- [0154] 도 6a를 참조하면, 감정 인식 수단의 훈련과정에서는 추출부에 의하여 추출된 특징 파라미터를 이용하여 감정 모델을 구축한다. 이렇게 복수의 감정에 따른 음향 모델을 구축할 수 있으며, 이는 메모리(1160)에 저장되어 사용자의 감정 판단에 사용될 수 있다. 이러한 감정 모델은 다수의 실험 참여자가 특정 상황을 간접 체험하는 동안 자연적으로 표출되는 감정을 이용한다.
- [0155] 본 발명에 따른 감정 모델의 구축은 실제 상황에서 자연스럽게 표출되는 음성과 유사한 음성을 사용하기 때문에 실용화에 유용하며, 화자 수가 많아 일반화에 유리하다. 이러한 본 발명에 따른 감정 인식 수단의 인식률 및 인식시간에 대한 실험 결과는 도 7과 같다. 도 7에서 볼 수 있듯이, 본 발명의 감정 인식 어플리케이션이 구비된 단말기는 매우 높은 감정 인식률을 나타낸다.
- [0156] 도 6b를 참조하면, 감정 인식 수단의 인식과정에서는 추출부에 의하여 추출된 특징 파라미터를 이용한다. 인식부는 입력 음성으로부터 추출한 특징 파라미터를 메모리(1160)에 저장된 감정별 음향 모델과의 비교를 통해 관측 확률값을 계산하며, 가장 높은 확률값을 나타내는 음향 모델에 대응하는 감정을 사용자의 감정으로 판단한다.
- [0157] 한편, 도 8은 본 발명의 감정 인식 수단이 사용자의 감정을 인식하는 과정을 개략적으로 나타낸 것이다.
- [0158] 도 8을 참조하면, 감정 인식 수단은 원음성으로부터 사용자의 음성을 감지하며, 상기 사용자 음성의 감지 과정에서는 신호 에너지가 이용될 수 있다. 사용자의 음성을 감지한 후, 특징 추출부는 MFCC와 Log-energy를 포함하는 특징 파라미터를 추출한다. 사용자 음성의 특징 파라미터는 메모리(1160)에 저장된 감정 별 음향 모델과 비교되며, 이 과정에서 획득한 인식 결과에 의하여 사용자의 감정을 정확하게 판단할 수 있다.
- [0159] 여기서 분류된 복수의 감정은 전술한 것과 같이, 중립, 기쁨, 분노, 슬픔 등이 될 수 있다. 단, 본 발명의 내용이 이에 한정되는 것은 아니고, 다른 기준에 따라 복수의 감정(예를 들어, 놀람, 혐오, 공포 등)으로 분류되는 것도 가능하다.
- [0160] 본 발명의 감정 인식 어플리케이션이 구비된 단말기는 사용자로부터 단말기(1100)에 내장 설치된 감정 인식 어플리케이션의 실행 명령을 입력받게 되면, 감정 인식 수단과 관련된 소정의 실행창이 디스플레이부(1151)에 표시된다. 상기 실행창은 디스플레이부(1151)에 팝업으로 나타날 수 있다. 사용자는 실행창을 통하여 사용자의 음성의 녹음과 관련된 녹음 정보를 입력하고, 마이크(1122)을 이용하여 사용자의 음성을 포함하는 제 1 음성을 녹음한다.
- [0161] 제 1 음성이 녹음되면, 감정 인식 수단으로 상기 제 1 음성이 전송되며, 감정 인식 수단은 메모리(1160)에 저장된 음향 모델 데이터와와 상기 제 1 음성을 이용하여 복수의 감정 중 사용자가 느끼고 있는 감정을 판단한다.

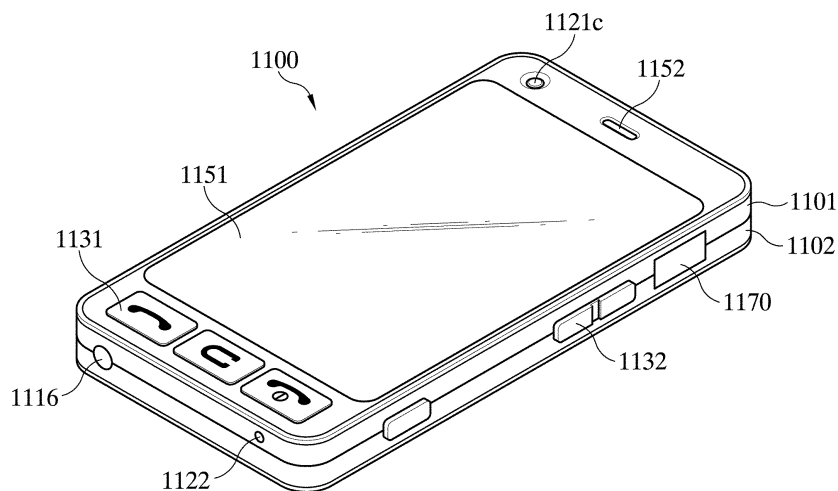
- [0162] 이러한 감정 인식 수단의 감정 인식 과정은, 상기 제 1 음성의 신호 에너지(signal energy)에 기초하여 상기 제 1 음성에 녹음된 상기 사용자의 음성을 감지하고, 상기 감지된 상기 사용자의 음성에서 제 1 특징 파라미터를 추출하며, 상기 제 1 특징 파라미터를 상기 음향 모델 데이터와 비교함으로써 수행될 수 있다.
- [0163] 이후, 출력부(1150)는 상기 제 1 음성에 기초하여 인식된 사용자의 감정에 대응하는 반응을 출력한다. 상기 반응은 디스플레이부(1151), 음향 출력 모듈(1152), 햅틱 모듈(1155), 방향부 등을 사용하여 이루어질 수 있다. 상기 디스플레이부(1151), 음향 출력 모듈(1152), 햅틱 모듈(1155) 등에 따른 출력은 메모리(1160)에 미리 저장된 특정한 패턴을 갖는 반응이다.
- [0164] 상기 출력부(1150)의 반응은 디스플레이부(1151)에서의 텍스트, 배경화면 또는 이모티콘 표시, 음향 출력 모듈(1152)에서의 설정된 음향 출력, 햅틱 모듈(1155)에서 소정 세기와 패턴을 갖는 진동의 발생, 방향부에서 발생되는 향기 등이 있다.
- [0165] 또한, 사용자가 본 발명에 따른 단말기(1100)를 이용하여 상대방과 통화 중인 경우에, 상기 단말기(1100)는 사용자의 감정을 인식할 수 있을 뿐만 아니라 상기 상대방의 감정을 인식할 수도 있다. 이 경우, 단말기(1100)의 출력부(1150)는 상기 인식된 상대방의 감정에 대응하여 적절한 반응을 출력할 수 있으며, 이에 따라 사용자는 통화 상대방의 감정을 더욱 잘 이해할 수 있다. 물론, 사용자는 자신의 감정상태에 따른 반응을 제공받을 수도 있으며, 이에 따라 사용자 자신의 감정의 완화에 도움을 받을 수도 있다.
- [0166] 한편, 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의해 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행할 수 있다. 그리고, 본 발명을 구현하기 위한 기능적인(functional) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있다.
- [0167] 또한, 상기와 같이 설명된 장치 및 방법은 상기 설명된 실시례들의 구성과 방법이 한정되게 적용될 수 있는 것이 아니라, 상기 실시례들은 다양한 변형이 이루어질 수 있도록 각 실시례들의 전부 또는 일부가 선택적으로 조합되어 구성될 수도 있다.

도면

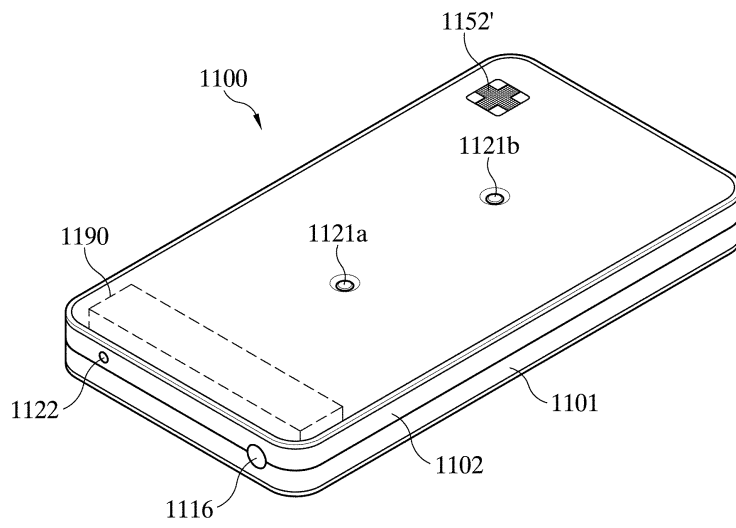
도면1



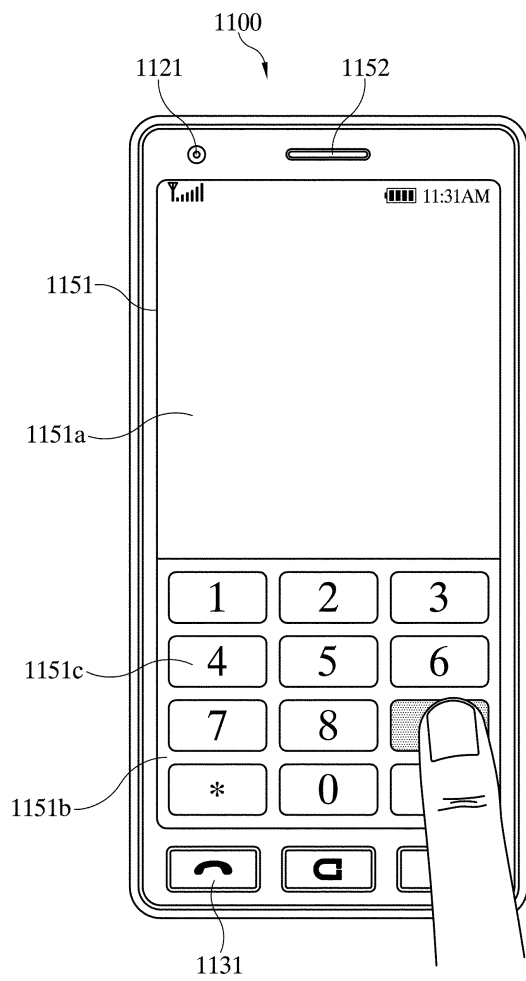
도면2a



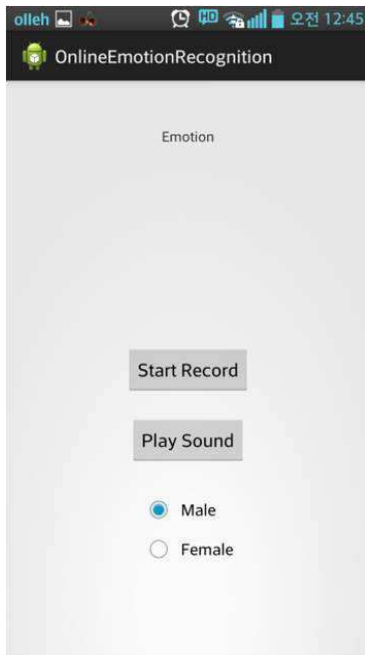
도면2b



도면3



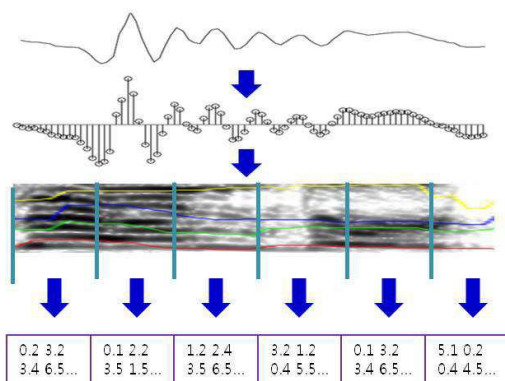
도면4



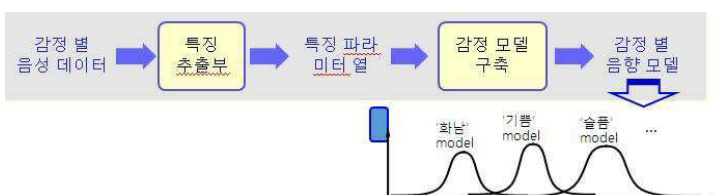
도면5a



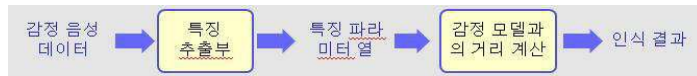
도면5b



도면6a



도면6b



도면7

감정 수	인식을 및 인식 소요 시간 (모바일 환경)	인식을 및 인식 소요 시간 (PC 환경)
2 (중립, 기쁨)	91%, 1초 (2초 길이 발화 기준)	97%, 1초
3 (중립, 기쁨, 분노)	75%, 1.3초	89%, 1초
4 (중립, 기쁨, 분노, 슬픔)	64%, 1.5초	86%, 1초

도면8

